

Paper for DISCOURSE STUDIES

MULTIMODAL ENACTMENT OF CHARACTERS IN
CONFERENCE PRESENTATIONS

Julia Valeiras-Jurado

Noelia Ruiz-Madrid

Version 3 (27/02/2019)

Julia.valeirasjurado@ugent.be

madrid@uji.es

Short title: MULTIMODAL ENACTMENT

Size in words: 8600

ABSTRACT

In academic oral genres such as conference presentations speakers resort to more than words to convey meaning. Research also suggests that persuasion, an important element of the communicative purpose of conference presentations, is frequently achieved through a combination of semiotic modes. Therefore, a skilful orchestration of these modes can be considered key to achieving effective communication in this genre.

However, our understanding of persuasion has often focused on specific elements of the message considered in isolation and mainly from the linguistic perspective. Relatively little attention has been paid to the overall persuasive effect achieved by the complex multimodal ensemble. This study approaches the analysis of persuasive strategies in conference presentations combining Multimodal Discourse Analysis and ethnographic methods. It focuses on a particular attention getting technique: enactment of characters, or acting the part of a person that is being referred to. Our analysis shows how it is achieved through the orchestration of different modes such as words, intonation, gestures, head movements, gaze and facial expression.

KEYWORDS: multimodal discourse analysis, academic discourse, conference presentations, attention getting, persuasion, strategies, enactment.

Bio notes:

Julia Valeiras-Jurado holds a Doctor's Degree in Linguistics from Jaume I University and Ghent University. She lectures business and scientific English at Ghent University. Her research interests are oral discourse and genres, multimodality and persuasive language.

Noelia Ruiz-Madrid is senior lecturer in the Department of English Studies at Universitat Jaume I. Her research focuses on discourse analysis, genre and integrating technology in the EFL classroom. Her more recent publications have appeared in *Ibérica* and *Multimodal Communication*, among others.

INTRODUCTION

The conference presentation (CP) is a widely used genre within academia. According to Fortanet-Gómez's (2005) classification, CPs fit into one of the three main categories of spoken academic genres, i.e.:

i) Classroom genres, including lectures, seminars, student's presentations and oral exams, among others;

ii) Institutional genres, which include academic year opening lectures, Honoris Causa speeches and Rector's addresses to the faculty, among others;

iii) Research genres, which can be further subdivided into a) *conference genres*, that include plenary lectures, poster presentations, workshops and the conference presentations that are central to this study and b) *other research genres* such as PhD thesis defences, Master's thesis presentations, etc.

Previous research has described the CP as part of a macro-generic event: the conference. Ventola (1999) and Forey and Feng (2016) use the term *semiotic spanning* to explain the connection between the CP and the discursive practices of the participants during the whole conference event. They argue that every presentation is influenced by the other presentations in the conference event and by related genres such as the abstract or the research article.

The CP has also been described as a process genre (Rowley-Jolivet and Carter-Thomas, 2005; Swales, 1990), because they often present research that is still ongoing. This dynamic research process tends to be diluted and idealized later in the written paper

(Hyland, 2009; Thompson, 2002). Research on CPs has a long trajectory. Dubois (1982) is an early study that presents a detailed account of biomedical conferences. Later Rowley-Jolivet (1999) claimed that CPs have a pivotal role in academic research, fulfilling three main functions: to present a scientific novelty, to give visibility to the research and to reinforce social cohesion within the discourse community. This line of research was continued by Carter-Thomas and Rowley-Jolivet (2003) and Rowley-Jolivet and Carter-Thomas (2005), who studied presentations from three different approaches:

- i) a microscopic bottom-up analysis to identify recurrent linguistic features and relate these choices to the communicative context;
- ii) a macroscopic top-down analysis (analysis of moves) to explore the rhetorical structure;
- iii) a multimodal approach that allows them to identify three main modes in CPs: language, visual communication and gesture.

In these studies, they highlight two important features of conference presentations that are of particular relevance for the present study: they are inherently persuasive and multimodal. We will now focus on these features in particular.

Valeiras-Jurado (2015) offers an overview of the literature on CPs that brings to the fore the multimodal nature of this genre and how it frequently involves, among other modes, the use of paralinguistic features (such as intonation, pauses, speed, etc.) and kinesic features (such as hand gestures, head movements, posture, etc.). For example,

Hood and Forey (2005) studied the use of gestures in introductions of plenary presentations, and found that they express attitudinal meaning: speakers use gestures to subtly evaluate their presentations in positive terms and encourage alignment between audience and presenter. Querol-Julián (2011) provides a detailed study of the multimodal expression of evaluation in discussion sessions of specialized conference paper presentations. She concludes that kinesic and paralinguistic features have an important role in the meaning-making process during the Q&A session of CPs. They have a pragmatic and modal function, showing the speaker's attitude or intensifying the evaluative meaning expressed in words. Along this line, Ruiz-Madrid and Fortanet-Gómez (2015) look into the role of non-verbal communication co-occurring with asides in conference plenary presentations. Results show that the analysis of linguistic, paralinguistic and kinesics features provides a more comprehensive understanding of the pragmatic dimension of asides in CPs in terms of function and communicative aim.

A particularly detailed multimodal study of the CP, including both kinesic and paralinguistic features is provided by Rendle-Short (2006). She shows how speakers indicate full engagement with the audience during topic talk through positioning, gaze and action. Topic talk is the section of talk that comes after the speaker has made a pause and reintroduced the topic at hand. During these sections of talk, the body usually faces the audience to show engagement, and speakers look at their audiences, sometimes intermittently as attention is also paid to visuals. A third way to show engagement is gesticulating with hands, in particular iconic and metaphoric gestures that visually represent part of the referential content, or beats that mark the rhythm of discourse and are usually synchronous with stresses. Another relevant finding of these studies is that

speakers often use kinesic and paralinguistic aspects to mark the macrostructure of their discourse, in particular to announce a topic transition or signal that a long pause is coming (so that it is not perceived by the audience as problematic).

To sum up the points made so far, previous research on CPs has proved that speakers fulfil their communicative aims through a skilfully orchestrated use of linguistic, paralinguistic and kinesic means. In particular, gestures, intonation and gaze serve as signposts to walk the audience through the presentation, making the message easy to understand and so contributing to fulfil the ideational function of construing human experience (Halliday, 1994). They also help achieve a desirable relationship with the audience, showing (dis)engagement and contributing to the expression of affect and evaluation, in this way contributing to fulfil the interpersonal function of facilitating human relationships as we communicate.

Therefore, speakers in CPs appeal to their audience by both ideational and interpersonal means. Proof of this is the important role that persuasion plays in this genre and which requires good use of both metafunctions. In CPs content needs to be interesting for the audience and relevant for the event, but it also has to be presented in a persuasive way. The notion that CPs are persuasive is shared by numerous authors. Räisänen (2002), for example, suggests that the function of conference presentations is to publicize, critically discuss and ratify research. Therefore, because their primary aim is to convince an audience of the validity, originality and usefulness of a piece of research, conference presentations can be considered to be a persuasive genre. Swales and Burke (2003), Wulf et al. (2009) or Querol-Julián (2011) also consider CPs as part of the

contingent repertoire of academic discourse. As Ruiz-Garrido (2015) points out, this persuasive communicative goal has important implications for the genre, because it requires the development of an interpersonal relationship with the audience who is physically present. Fernandez-Polo (2014: 166) summarizes these views when he says that

[m]odern scientific communication is strongly argumentative and persuasive. In the highly competitive context of present-day science, stressing and persuading others of the value of our claims has become crucial, and consequently the building up in the text of favourable personal relationships with the audience.

An example of this trend towards persuasion in oral academic discourse is also shown by Thompson (2002), who claims that speakers in CPs resort to what she calls ‘involvement strategies’ such as *storytelling* and *constructed dialogues*, to persuade an audience to accept their ideas. She points out how storytelling has the effect of engaging the listener. She demonstrates how narratives in past tense during conference presentations relate difficulties and failures during the research process. Sometimes these narratives include self-deprecating humour and are spiced up with direct speech, colloquialisms and idiomatic expressions. In contrast, in the written paper these stories become plain present tense results. The use of narratives is a persuasive appeal that helps create rapport: it shows the person behind the study and provides credibility. This was already noted by Ochs et al. (1994), who argue that dynamic grammar and gesturing help physicists create a visual representation of experiments so that they are perceived as credible. Morton (2009) also found similar attempts to increase vividness in student architecture presentations through gesture, language, visuals and physical objects. Taken

a step further, this vivid account of research practices can even take the form of enactments that are performed multimodally, as we will illustrate in this paper.

We understand that a (re)enactment takes place when the speakers act the part of a person that is being referred to and do not limit themselves to narrating previous events using reported speech, but they rather represent these events using other embodied modes such as gestures, intonation and facial expression. This is a tendency noted in naturally occurring conversation and reported by Sidnell (2006), Good (2015) and recently by Soulaïmani (2018). They point out that these (re)enactments prompt a higher level of attention, reciprocity and alignment from the addressee, which explains why in the context of a presentation they can work as an involvement strategy and a persuasive appeal for the audience, as noted in the previous paragraph.

Although narratives and (re)enactments are a proven persuasive appeal and a powerful tool to engage the audience, they are also a challenge for novice presenters, who frequently struggle to present their research in a persuasive way. The challenge becomes greater when we take into account that these presentation skills require the mastery of a whole range of semiotic modes. Most training programs lack a systematic multimodal perspective, and a metalanguage that can help identify and effectively use all the semiotic modes available to convey persuasion when giving presentations (Forey and Feng, 2016). We believe that a deeper understanding of multimodal persuasive strategies such as attention getting through enactment of characters can be a great step forward in this sense.

METHOD

In this paper we present a video-based analysis of persuasive strategies (Valeiras-Jurado, 2015; Valeiras-Jurado and Ruiz-Madrid, 2015; Valeiras-Jurado et al., 2018) in a corpus of CPs, with focus on one specific strategy: attention getting through enactment of characters. The analysis combines MDA and ethnographic methods: interviews and observations. A more detailed account of this methodology can be found in Valeiras-Jurado et al (2018).

Data collection

The event selected for data collection was an international symposium on business discourse. The title of the event was *The Ins and Outs of Professional Discourse*, and special attention was paid to two main topics: access to business environments (input) and feeding back research results to business communities (output). The event took place in a large room equipped with a small podium, a lectern with a microphone attached, a projector, a screen and a computer. The audience ranged from 14 to 23 persons. Each conference presentation was allotted 40 minutes, including a Q&A session.

Consent was obtained from the speakers to video-record and analyse their presentations. The presentations were video-taped using a fixed camera that focused on the speaker. A total of 13 presentations were recorded and from these, 5 were selected on the basis of the following criteria:

- i) Good quality of image and sound of the recording;
- ii) Availability of the complete presentation in the recording;

iii) Availability of speakers for follow-up.

During the recordings, both researchers were present and used observation sheets to take notes providing contextual information regarding the event (e.g. date, organisers, frequency, etc.); the presentation (e.g. title, length, order in the program); the speaker (e.g. name, age, gender, mother tongue, performance) and the physical environment (e.g. the room, the audience, devices, etc.).

Two different types of interviews were carried out during the study: a semi-structured interview after each presentation and an open interview at a later stage to discuss the results. The first interview was a face to face semi-structured interview about aspects such as:

i) The speakers' motivation to participate in the event, which ultimately determines their main goal in their presentations, something crucial to fully understand the communicative intentions lying behind their use of modes.

ii) What they knew about the event beforehand (e.g. size and type of audience, topics expected, types of presentations expected, etc.) and the way they prepared accordingly, which is also important to fully understand how speakers were using the modes, and in particular to determine how much of this was spontaneous and how much the result of practice or explicit training.

iii) Their satisfaction with their performance. This information proved useful to focus on particularly persuasive moments.

iv) Their opinion of other presentations as members of the audience.

The second interviews took place at a later stage and were used to discuss results with speakers and cross-check interpretations. Details about these interviews are provided in the section ‘Triangulating results’. For the sake of consistency, the same researcher carried out all the interviews. The language used was English, they were audio recorded and the researcher took written notes.

The information provided by the observation sheets and the interviews proved crucial, first in the selection of the data to be analysed, and later on in its interpretation, since the data provided by the interviews helped interpret the speakers’ intended communicative effect and their use of persuasion.

The corpus

The total size of the corpus is 173 minutes. The size of the corpus does not allow for quantitatively-based generalisations, but it is valid for qualitative analysis. It is also in line with previous multimodal studies, which due to their minute level of detail and the lack of automating tools cannot afford the use larger corpora (Querol-Julián, 2011; Morell, 2015). The total number of speakers is 7, since 2 of the presentations were co-presented. All the speakers are expert users of English as a Lingua Franca (ELF) in their academic activity, regardless of whether or not they were native English speakers (Graddol, 2003; Jenkins, 2011). Regarding expertise in presenting at conferences, only one of the native speakers (co-presenting) reported being novice (one previous presentation experience). The delivery style of the speakers was classified according to Dudley-Evans and Johns (1981) by taking into account how much speakers resort to their

notes: reading style means speakers read through their notes, conversational style means speakers follow their notes without reading through them and rhetorical style means speakers rarely employ notes. All of the speakers in the corpus used a conversational style. Table 1 provides an overview of the corpus of CPs.

Table 1: The corpus

	Gender	English proficiency	Expertise	Delivery style	Support/ devices	Duration
CP1	Female	Expert user	High	Conversational	PPT Lectern Microphone	00:28:44
CP2	Female	Expert user	High	Conversational	PPT Lectern Microphone Written notes	00:35:20
CP3	Sp 1 female Sp 2 female	Sp 1 Expert user Sp 2 Expert user	Sp 1 High Sp2 Novice	Conversational	PPT Lectern Microphone Written notes Sp 1: Handouts	00:25:05
CP4	Female	Expert user	High	Conversational	PPT Lectern	00:43:54

					Microphone	
CP5	Sp 1 Female Sp 2 Male	Sp 1 Expert user Sp 2 Expert user	High	Conversational	PPT Lectern Microphone Sp 2: pointer	00:40:53

Rich points for analysis

From the corpus of CPs a series of potentially persuasive points were selected for analysis, which we call *rich points* (Valeiras-Jurado, 2015). The term *rich point* was coined by Agar (1996). It refers to moments of experience that stand out from the rest because they reveal a cultural difference, which makes them worthy of attention as a research focus. We borrow this term to refer to moments in the presentation in which the speakers are focusing on persuading the audience and therefore are applying different persuasive strategies. Their position within the presentation was not taken into account during selection, and therefore varies in each case. Their length also varies, since the contours were established according to a sense of completeness: they seem to fulfil a communicative intention or objective (e.g. illustrate a claim, prompt the audience to action, etc.). For the purpose of keeping the analysis within feasible limits we restricted the number of rich points to two in each presentation (the ones that stand out as more persuasive in relation to the rest of the presentation).

The identification, delimitation and selection of rich points were done using an *inductive* approach to video data (Goldman et al., 2007). The selection process was first tested by a pilot analysis in which two researchers viewed the same presentation and independently selected two rich points using the same criterion: high persuasive effort from the speaker. The researchers obtained the same results, indicating that the perception of high persuasive efforts is not (entirely) subjective. The information gathered from the observation sheets and the first interviews with speakers also contributed to double check the selection process. For example, if the rich point selected coincides with a moment the speakers felt particularly satisfied with, we consider this as corroboration of our selection.

The use of rich points was adopted to avoid prioritizing any semiotic mode and to keep the focus on the multimodal ensemble (Kress, 2010). Instead of using one of the modes as the driver of the analysis, we are more interested in elucidating how the different modes interact to create a persuasive message (Valeiras-Jurado, 2015; Valeiras-Jurado et al., 2018). Table 2 offers an overview of the rich points selected. The total size of the conference presentations rich points is 7.2 minutes. To keep track of these rich points while preserving the anonymity of speakers, they were coded according to the genre (CP), the initials of the speakers and then numbered.

Table 2: Rich Points

CONFERENCE PRESENTATIONS RICH POINTS			
CODE	Begin	End	Duration
CPDO1	00:05:14	00:06:00	00:00:46
CPDO2	00:17:00	00:17:24	00:00:24

CODE	Begin	End	Duration
CPAS1	00:16:05	00:16:35	00:00:30
CPTO1	00:19:26	00:20:13	00:00:48
CODE	Begin	End	Duration
CPAM1	00:03:13	00:03:37	00:00:24
CPAM2	00:21:10	00:21:58	00:00:48
CODE	Begin	End	Duration
CPRE1	00:00:38	00:01:36	00:00:58
CPKE1	00:15:24	00:16:14	00:00:48
CODE	Begin	End	Duration
CPPE1	00:04:43	00:05:22	00:00:39
CPPE2	00:07:07	00:08:12	00:01:05

It must be stressed that the rich points were not selected because they illustrated a particular linguistic, paralinguistic or kinesic feature, but exclusively on the grounds of strong persuasive effort, regardless of the modes employed. In the remaining of this paper we will focus on three rich points that make use of a particular attention getting technique: enactment of characters or acting the part of a person that is being referred to.

The semiotic modes and the instruments for the Multimodal Discourse

Analysis

A multimodal discourse analysis (MDA) of the rich points was performed with the help of specialised software. This MDA entailed the use of annotated transcripts including data regarding all the modes that integrate the analysis: words, intonation, gestures, head movements, gaze and facial expression. The rationale for selecting these modes is that they are always there in any instance of oral discourse and they deeply

influence the way a message is understood. A brief overview of our approach to each mode follows in the next paragraphs.

We use the term *words* to include speakers' choices in terms of *lexis*, *grammar*, *style* and *register*. These choices include evaluative language, three-part lists, symmetrical syntactic structures, inclusive pronouns, examples, comparisons, narratives, among others (Carter, 1997; Hyland, 2009; Lakoff, 1982). We believe that they can have an effect on persuasion, because they help provide the message with characteristics that, according to literature, make the text persuasive (e.g. they make the message easy to understand and remember). We use the term *words* because we consider that it describes this mode more accurately. The term "verbal mode" can be misleading in the context of the present study, because it can also include material that is produced verbally (i.e. using the articulatory organs) but which is non-linguistic, and therefore would fall into "paralanguage". The term "lexical" would be too restrictive, because under the semiotic mode *words* we also consider aspects such as grammatical structure, which is usually considered out of the scope of *lexis*. Finally, the term "language" would be far too generic, because language can be considered a macro-system including paralinguistic and kinesic features (Poyatos, 1983).

Intonation is frequently defined as variations in pitch that make up a pitch contour. Our approach to intonation is *Discourse Intonation (DI)*. The main tenet of this approach is that speakers make meaningful choices in their use of intonation that reveal their assumptions about the interaction in the communicative process (Brazil, 1997). Discourse intonation divides speech into *tone units*, each one containing one or two

prominent syllables. Prominent syllables are louder and longer. In our transcripts they are capitalized. The first prominent syllable in a tone unit is called *onset*, the second is called *tonic* syllable. *Tone* is the pitch movement that begins in the tonic syllable and continues throughout the tone unit. In our transcripts we represent tone with arrows. *Key* refers to the relative pitch of onset syllables in relation to the onset syllable of the previous tone unit, while *termination* refers to the relative pitch of tonic syllables in relation to the onset syllable in the same tone unit. In our transcripts a high key and termination is represented as superscript, and low key and termination as subscript. The following example is a tone unit pronounced with rising tone and high termination:

i SEEMED to be the only person that ^{WASN'T} a psychiatrist there ↗

Kendon (2004: 7) defines *gesture* as ‘visible action when it is used as an utterance or as part of an utterance’. Following Kendon, we limit our concept of gestures to hand and arm movements. We focus on the type and function of the gesture drawing on Bavelas et al. (1995), Kendon (2004), McNeill (1992) and Querol-Julián (2011). We distinguish between the following types: 1) ‘*iconic*’ (represent concrete objects and events), 2) ‘*metaphoric*’ (represent abstract ideas), 3) ‘*beats*’ (repetitive gestures that usually mark the discourse flow) and 4) ‘*deictic*’ (point to something). We also differentiate the following functions: 1) ‘*referential*’ (they are part of the referential content), 2) ‘*pragmatic*’ (they show the attitude of the speaker towards the content and indicate how content is to be interpreted), 3) ‘*interpersonal*’ (they regulate interaction) and 4) ‘*cohesive*’ gestures (they connect thematically related but temporally separated parts of discourse). We also draw on Kendon’s gesture families, i.e., gestures with similar

kinesic characteristics that seem to share a common semantic theme, with each gesture within the family adding a particular semantic nuance. For example the ‘open *hand supine*’ (OHS) family includes gestures made with the palm of the hand facing up, and share the semantic theme of offering or willingness to receive something.

Regarding *head movements* we focus on nods, shakes and lateral movements. We consider it interesting to look at the amplitude of the movement and the number of repetitions because they can provide insights into pragmatic meaning. For example Hadar et al. (1983) note how a nod can have different meanings depending on the number of repetitions (one nod means polite involvement, two nods real interest, but three nods can convey impatience).

In the case of *gaze*, we distinguish whether the speaker is looking at the audience, the screen, the lectern, or keeps eyes closed, with a decreasing degree of audience engagement in each case (Forey and Feng, 2016).

For *facial expressions* we focus on what Knapp and Hall (1992) call *semantic displays*: facial actions that are connected with the content of what is being said and can be redundant or complementary to the rest of the modes.

The software that supported the MDA consisted of the programs PRAAT and ELAN. The tool for phonetic analysis PRAATⁱ allowed us to obtain accurate measurements of pitch and intensity. With the annotation software ELANⁱⁱ we were able to transcribe and annotate audio and video files. Transcriptions and annotations are then organised on tiers as shown in Figure 1.

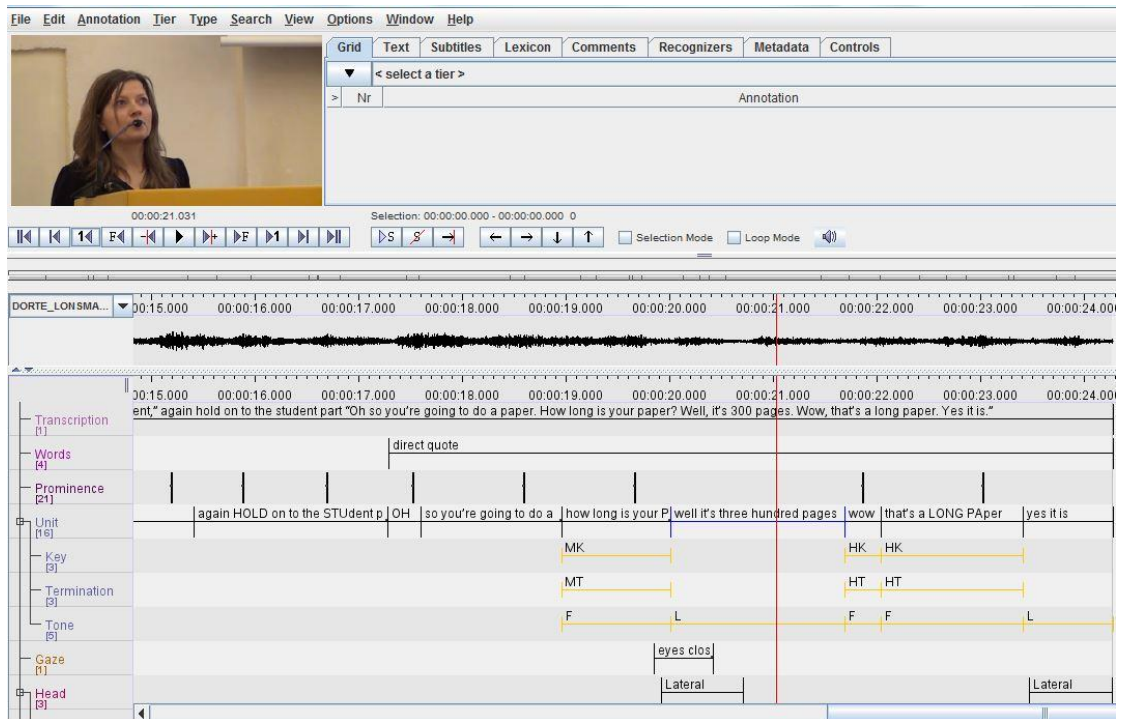


Figure 1. Annotation with ELAN

Further details about our use of tiers, transcriptions and annotations can be found in Valeiras-Jurado et al. (2018).

Triangulating results

Ethnographic interviews were used to discuss preliminary results with speakers and cross-check interpretations. This exchange of interpretations provided the analysis with more reliability. Occasionally, speakers offered alternative interpretations, which were taken into account and integrated in our discussion.

Four of these interviews took place face to face and three were online video calls. The time span between these interviews and the presentations was considerable, because they necessarily needed to take place once our MDA analysis was completed. Because of

this time span we considered that the most appropriate approach was an open interview including a stimulated recall with videos of the rich points. The interviews started with an explanation of the content and purpose of the interview. Then the researchers visualised the rich points with the speakers, without revealing results or annotations. The visualisation of the rich points allowed us to discuss aspects of the multimodal behaviour of the speakers that the analysis had revealed as relevant for the persuasive effect (e.g. a particular use of intonation or a specific gesture). In addition, with this methodology we were able to prompt the speakers' interpretations before we shared our own in the final stage of the interview.

RESULTS AND DISCUSSION


In this paper we present a multimodal analysis of three rich points in CPs that feature the use of a particular persuasive strategy: *attention getting*. This strategy is used to raise and maintain the interest of the audience. Andeweg et al. (1998) use the term *exordial techniques* to refer to this concept and it is also popularly known as 'hook' or 'attention-getters'. It can take a variety of forms (i.e. realised by means of a variety of semiotic modes) including a particularly marked use of intonation (e.g. a high pitch), a marked use of gesture (e.g. wider amplitude or more repetitions) or the use of stylistic devices such as narratives or direct addresses to the audience. Attention getting techniques are especially relevant in the openings of the presentations, but can also appear later on.

For the purposes of the present study we zoom in on one particular type of attention getting: enactment of characters. In these cases the speakers act out assuming

the roles of people they are referring to in their presentations, such as participants in their research or other researchers. We consider it a way of attention getting, since it turns an account of research process into something more vivid and potentially more likely to get the attention of the audience. This is the case in the following example (Rich point CPKE1), in which the speaker is relating how she tried to disseminate her research in a previous conference. Table 3 offers the transcript corresponding to this excerpt.

Table 3: Example of narrative

DI (Discourse Intonation) transcriptionⁱⁱⁱ	
1.	so the Apa annual meeting was the BEST opportunity i could really find
2.	to disSEminate my reSEARCH
3.	eh
4.	there are THOUSands of PEOple there
5.	the WHOLE
6.	eh
7.	occupational community GAtchers to
8.	eh
9.	listen to TALKS
10.	it's LARGEly
11.	eh
12.	like THIS
13.	but MORE for the psyCHIAtric community
14.	i presented a POSter
15.	in young investigator's POSter session
16.	and eh
17.	THAT'S my poster right there

18. it was ACTually INTEResting because
19. i SEEMED to be the only person that ^{WASN'T} a psychiatrist (1) there 
20. and a LOT of people would WALK by
21. and they would DOUble-take and look (2)
22. like almost that i didn't beLONG there
23. but ONCE i was able to TALK to them about it
24. a lot of people were VErY interested
25. on the IMPact
26. that RHEtoric has on psyCHIAtry
27. and how they could USE that PRACtically

Concerning the use of words as a semiotic mode, the speaker uses the first person pronoun 'I' to present herself as the protagonist of a first-person narrative. It is also interesting that to contextualise this narrative the speaker makes a reference to the event where she is presenting now and compares it to the event she is referring to in her narrative. This connects the speaker's story to the personal experience of the audience, and can be considered both an attempt to build rapport and an attempt to fit in the larger event of the conference, a common trend in academic discourse spotted by Hyland (2009). It equally provides an example of what Ventola (1999) would call *semiotic spanning*.

Regarding intonation, high termination in tone unit 19 (see transcript above) adds a sense of surprise (Brazil, 1997) and represents the reaction of the people seeing her poster. The speaker is not using the exact words of the passing psychiatrists, but she is mimicking their intonation to vividly convey their surprise. At the same time a very marked rise tone projects the communicative dominance of more experienced researchers

questioning the validity of her research (Brazil, 1997). A micro pause before ‘there’ in tone unit 19 emphasizes the deictic reference to the APA event she is describing. There is a two-second pause after unit 21, which is filled by a facial expression that mimics the reaction of surprise of the participants in the narrative that she is creating about her experience of presenting a poster (see Figure 5 below). We believe this is the moment in which this narrative becomes an enactment of characters, as we will argue later in this section. The prominences in units 23-27 emphasize the final positive reaction towards her research (after initial surprise and lack of understanding).

Turning now to gestures and head movements, the extreme quantification ‘thousands of people’ as she describes the APA event is accompanied by a head shake (Figure 2) that can be interpreted as denying any potential counter argument (Kendon, 2004): the speaker seems to be anticipating resistance from the audience regarding this extreme quantification (e.g. ‘thousands of people’ provoking the reaction ‘no, it can’t be that big’) and shakes her head to negate this counter-reactions. It is a pragmatic gesture that makes her narrative more vivid.



Figure 2. Head shakes in CPKE1

The reference to the event where the speaker is currently presenting in ‘largely like this’ is clarified by an open hand supine (OHS) deictic gesture (Kendon, 2004) pointing to the audience (Figure 3a). The gesture is held throughout units 10-13, and it is combined with a series of 5 beats. The first beat is synchronous with the prominence in ‘largely’, the second one is synchronous with the filler (eh), the third one precedes the prominence in ‘more’, the fourth coincides with this prominence and the last one precedes the prominence in ‘psychiatric’ (Figure 3 b). This combination of gestures (OHS deictic + beats) is partly referential, because it clarifies what ‘this’ refers to, and partly pragmatic in its inclusive effect, because it is also trying to make the audience feel identified with the setting of her narrative (i.e. a conference like the one you are attending now).

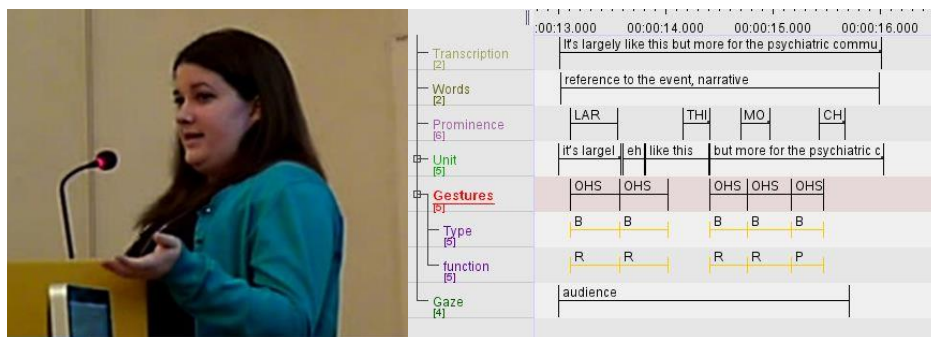


Figure 3a. Open hand deictic gesture in CPKE1. Figure 4b. Synchrony of deictics and beats in CPKE1

Towards the end of the excerpt, two closed hand beats moving downwards synchronous with ‘very interested’, emphasize the extreme evaluation (Figures 4a and 4b).

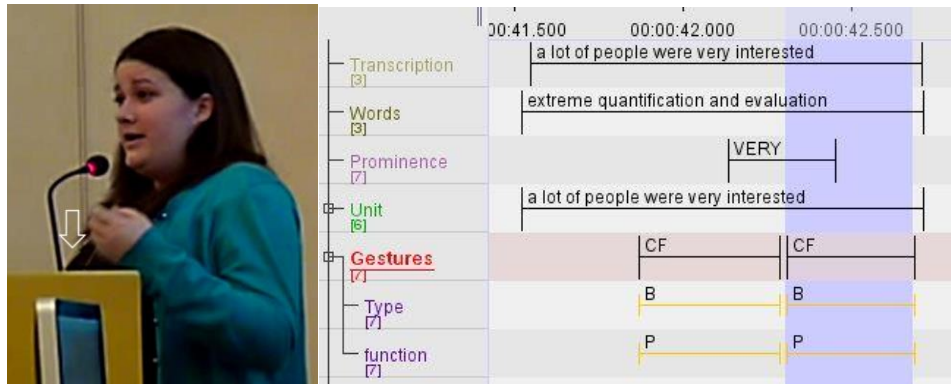


Figure 4a. Beats in CPKE1.

Figure 4b. Synchrony of beats in CPKE1

Concerning gaze, the speaker directs gaze mainly to the lectern at the beginning of the rich point, but interestingly starts looking at the audience as she makes a reference to the event ('largely like this') to get them engaged, adding to the effect of the OHS gesture discussed above. Later she directs gaze mainly to the screen, first to invite the audience to pay attention to the projection of her poster (an element in her narrative), and then to impersonate a character in her narrative: a participant in the conference who is watching her poster (Figures 5 and 6 below).

In what concerns facial expression, we believe it is interesting how the speaker smiles (Figure 5) as she pronounces 'that's my poster' (unit 17) and 'wasn't a psychiatrist there' (unit 19). The speaker is showing stance through facial expression (Soulaïmani, 2018), and inviting the audience to agree with this stance (i.e. prompting the audience to find the situation funny, as she did).



Figure 5. Smile in CPKE1

Finally, it is worth noticing how this speaker mimics the facial expression of the character in her narrative mentioned above. She adopts an expression of surprise, misunderstanding or disbelief with frowned forehead and pursed lips (Knapp and Hall, 1992), which is the way this person felt when seeing the poster she presented (Figure 6).



Figure 6. Face expression in CPKE1

In sum, the speaker is multimodally building an interesting narrative of her previous conference experience using words, intonation, head movements, gestures, gaze and facial expression in a multimodal ensemble (Kress, 2010). This can be considered an example of what Thompson (2002) calls an *involvement strategy through storytelling*,

and is also a way in which the speaker expresses emotions (i.e. how she felt when disseminating her research) to trigger the audience's empathy (Forey and Feng, 2016). Furthermore, the strategy is in line with the communicative intentions that the speaker reported in the first interview. She explained that her main goal was to stress the 'value of professional discourse research in occupational communities' because it was 'very understudied'. Thus, the vivid representation of how she disseminated this research and how the target community reacted to it reveals itself as particularly appropriate to achieve her communicative goal.

What we find really interesting in this example is that at one point in this narrative the speaker briefly plays the role of one of the characters (a participant in the conference where she disseminated her research) and mimics the facial expression this person adopted when seeing the poster she presented. In this way, the speaker goes beyond merely storytelling, and 'acts out' a character in her story. This example shows how speakers in CPs can provide 'another turn of the screw' in their narratives by adding a new semiotic mode to the ensemble (in this case facial expression) and turn their narratives into enactments.

The next example, CPDO2, shows how another speaker in one of the CPs gets the attention of the audience by impersonating a character in her narrative about her research experience doing ethnographic field work. In particular, she discusses how she temporarily assumes the role of a student to gain access to her informants and uses a particular example to illustrate the relationship informant-ethnographer. This is very much in line with what the speaker reported was the main point in her presentation during

the first interview: ‘we need to be aware of the different roles and positions as researchers’. She also mentioned that in order to do that it is useful to rely on ‘anecdotes from the field’, which is actually what she is doing in this excerpt. It is therefore plausible to think that the speaker is trying to make this excerpt particularly persuasive. In this case the enactment is more extensive. Table 4 shows the transcript of this rich point.

Table 4: Example 1 of enactment of characters

DI transcription	
1.	and when i showed up at FOUR thirty in the morning
2.	to FOLlow them for their day of WORK
3.	they were CLEARly apprehENSive about
4.	what i was DOING there
5.	was i trying to SPY on them cleaning the OFFices
6.	or WHAT
7.	what was the PURpose of my BEING there
8.	and i said i was PhD STUdent
9.	and they CAUGHT on to the STUdent part
10.	OH
11.	so you’re going to do a PAper
12.	how long is your PAper
13.	well it’s three hundred pages
14.	wow
15.	that’s a LONG Paper

16. yes it is

The use of direct speech, first person pronouns and informal speech (e.g. the phrasal verb ‘showed up’) makes the narrative more vivid and personal. The speaker marks the change of personae by using an overall higher pitch when assuming the role of the informants in her ethnographic research when she is not using her own words, but those of her informants (units 10, 11, 12, 14 and 15). This finding is congruent with Soulaïmani (2018), who also found that speakers use voice qualities like pitch to signal a shift in personae during re-enactments. This higher pitch at the same time conveys that the informants are surprised by the information they are receiving (Brazil, 1997). This is supported by the vocalizations in tone units 10 and 14, which are stereotypical ways of expressing surprise.

At the same time, she visually represents the lack of understanding of her informants through a sequence of subtle head shakes synchronous with the words ‘or what’, which adds to the vividness of the narrative (Figure 7). The speaker is not only reporting the feelings of her informants through words (i.e. they were apprehensive because they didn’t understand), but she is also visually representing this confusion through a head movement that can be considered an emblem for ‘no’, and frequently accompanies negations as in ‘I don’t understand’ (Kendon, 2002).



Figure 7. Head shakes in CPDO1

The speaker confirmed this interpretation during the second interview, but also added an additional one: ‘maybe I am trying to illustrate that the idea of me being a spy is a bit ridiculous, or maybe illustrating their confusion, like they don’t know what I’m doing there’.

The lateral movements of the head (Figure 8) visually represent the researcher’s feelings at the reaction of her informants.



Figure 8. Lateral head movements in CPDO1

In combination with a flat unenthusiastic intonation, or *oblique* in Brazil’s (1997) terms, and lack of eye contact (her eyes are briefly closed as shown in Figure 8), they seem to convey a sense of boredom at having to repeat the same thing over and over again. During the interview the speaker offered a different interpretation: it is a way of

minimising what she considers a ‘white lie’ in saying to her informants that she was a student. She did this to gain easier access during her ethnographic research and encourage these people to talk to her, but she acknowledges it was not totally accurate.

In this rich point we see four modes working together to create an enactment of a research experience. The screenshot of ELAN presented earlier in Figure 1 provides more details about the synchronicity of the four modes at play, and shows how the speaker is using a direct quote, level intonation, closed eyes and a lateral head movement simultaneously. When asked about her communicative intention the speaker reported that she was hoping to make it interesting for the audience:

I was trying to tell a story to try to kind of put them, yeah, in my place. So it was supposed to make it interesting, to spice things up a little. And again, it makes it personal, because I'm telling the story of what happened to me.

This example clearly supports Thompson’s (2002) arguments regarding the differences that can be found when comparing the methodology and results sections of oral presentations (vivid account of work in progress) with those of written papers (impersonal, plain present tense results). It is also in line with previous research that emphasizes the persuasive nature of CPs (Carter Thomas and Rowley-Jolivet , 2003; Rowley-Jolivet and Carter Thomas, 2005; Ruiz-Garrido, 2015), because it shows how the speaker feels the need to ‘make it interesting’ for the audience to find her presentation convincing.

In the following example, CPAS1, the speaker is discussing how during an internship program where she was involved as a teacher, the students and the company

did not manage to build the relationship of mutuality and trust that was required for the project, and this was evident in some communication gaps. It is interesting to note that during the first interview the speaker summarized the main point of her presentation as ‘the tension between the expectations of students and the professionals’, which largely corresponds to the content of this rich point and explains why the speaker would take particular care in making it persuasive. Table 5 shows the use of words and intonation.

Table 5: Example 2 of enactment of characters

DI transcription	
1.	eh
2.	also
3.	coming BACK to that concept of FREedom
4.	their
5.	one of their CORE principles
6.	the ^{STU} dents
7.	to ^{THEM}
8.	were VEry FREE
9.	eh
10.	when
11.	reMEMber when the STUdents said
12.	we weren't aLLOWED to interview eh emploYEES
13.	well in FACT
14.	it WASN'T because they weren't aLLOWED
15.	it just did not seem ^{RE} levant to the COMpany
16.	but in FACT all the resources were THERE
17.	eh

- | | |
|-----|---------------------|
| 18. | just |
| 19. | eh |
| 20. | if they WANTED them |
| 21. | it was it was THERE |

The use of the pronoun ‘we’ to refer to students turns this passage into a direct quote (it is the students talking) and provides a narrative style. The repetition of the adversatives ‘in fact’ and ‘but’ emphasizes the contrast between what students perceived and what the company perceived, and highlights the misunderstanding this passage refers to.

Concerning intonation, high key in ‘students’ (tone unit 6) and ‘them’ (tone unit 7), and later in ‘relevant’ (tone unit 15) conveys surprise (Brazil, 1997) and again highlights the contrast between what the company believed and what the students believed.

Turning now to head movements, repeated head shakes simultaneous with ‘did not seem relevant to the company’ contradict visually the students’ wrong interpretation (Figure 9).



Figure 9. Head shakes in CPAS1

The speaker is assuming the role of the company at this moment and representing their reaction, paraphrasable as ‘it’s not that we don’t allow you, we thought it wasn’t useful for you’. Note how the facial expression with frowned forehead also supports head movements in conveying lack of understanding (Knapp and Hall, 1992). The speaker agreed with this interpretation during the second interview and even intuitively hinted at the strategy of enacting characters saying that ‘as I was presenting my opinion about those tensions you sort of naturally take on the different roles’. Later she also added that this enactment of characters is actually very coherent with the main topic of her presentation: different roles in the internship program. In addition, she also explained that ‘when I see they [the audience] are engaged it makes me feel more comfortable’, which accounts for her efforts at getting the attention of the audience.

In short, in these examples speakers retain the attention of the audience using words intonation, gestures, head movements, gaze and facial expression to build a vivid narrative of what happened, and even providing a (partial) re-enactment of the episode.

Therefore, they can be considered multimodal instances of Thompson's (2002) involvement strategies: *story-telling* and *constructed dialogues*.

CONCLUSIONS

The three examples presented in the previous section show how the speakers in these CPs try to get the attention of their audience in order to create a convincing message. Furthermore, they do this orchestrating a variety of semiotic modes into persuasive multimodal ensembles. We have focused on a specific type of attention getting: enactment of characters. This enactment of characters can be seen as a way of taking narratives a step further. Speakers in these cases do not simply report what happened, but they re-enact it as if it was happening at that moment. This adds extra vividness and makes their accounts of research experiences more interesting, and in turn more likely to be accepted.

From a methodological point of view, we believe there is added value in combining MDA with ethnographic methods. On the one hand it enables us to focus on the multimodal ensemble without giving priority to any of the modes. If one mode is considered to be playing the leading role by default, the study of the other modes is restricted to examples of co-expression with that leading mode. On the other hand, this combined methodology gives us the possibility of triangulating results. A study of persuasion necessarily entails probing into the communicative intentions of the speakers, and this would have been difficult without discussing our results with the speakers. In all cases we prompted the speakers' interpretation before we offered them our own. Very

frequently it coincided with ours, but in some cases they offered a different point of view, which highly enriched our analysis.

The study presented in this paper focuses on certain modes in particular: words, intonation, gestures, head movements, gaze and facial expression. We are aware that these modes do not constitute the entirety of the complex multimodal ensemble that is orchestrated in oral discourse in general, and in CPs in particular. Furthermore, our results indicate that sometimes modes are difficult to delimit and it is not always possible to exclude them from the analysis, because they are so intertwined that their meaning contribution is interdependent.

We believe that a fascinating topic for further research would be to expand the scope of the multimodal ensembles studied to include more semiotic modes. Likewise, larger-scale, generic studies including different persuasive oral genres can allow for more representativeness, and might help to determine if certain multimodally realised persuasive strategies can be considered as characteristic of a particular genre (or of the changes a particular genre is undergoing). This knowledge can enrich our understanding of these genres, and can ultimately be used to improve materials for language training that use them as didactic tools.

Finally, it would be equally interesting to include the other side of the coin in the study: the reaction of the audience to the speakers' persuasive efforts (i.e. whether they make the presentation memorable, elicit interest, laughter, etc.). Along this line, works such as Ruiz-Madrid and Fortanet-Gómez (2015) show, for instance, how speakers convey intentional humour through multimodal ensembles as they make autobiographic

references, in order to keep the attention of the audience and contribute to a relaxed atmosphere.

REFERENCES

- Agar, M. (1996) *The Professional Stranger: An Informal Introduction to Ethnography*. London: Academic Press.
- Andeweg, B. A., de Jong, J. C. and Hoeken, H. (1998) 'May I have your attention? : Exordial techniques in informative oral presentations', *Technical Communication Quarterly* 7(3): 271-284.
- Bavelas, J., Chovil, N., Coates, L. and Roe, L. (1995) 'Gestures specialized for dialogue', *Personality and Social Psychology Bulletin* 21(4): 394-405.
- Brazil, D. (1997) *The communicative value of intonation in English* (2nd ed). Cambridge: Cambridge University Press.
- Carter, R. (1997) *Working With Texts: A Core Book for Language Analysis*. London: Routledge.
- Carter-Thomas, S and Rowley-Jolivet, E. (2003) 'Analysing the Scientific Conference Presentation (CP): A Methodological Overview of a Multimodal Genre', *ASp: La Revue Du GERAS* 40(39-40): 59-72.
- Dubois, B. L. (1982) 'And the last slide please: Regulatory function at biomedical meetings', *World Language English* 1(2): 263-268.
- Dudley-Evans, T. and Johns, T. F. (1981) 'A team teaching approach to lecture comprehension for overseas students'. In: Dudley-Evans, T. and Johns, T.F. (eds) *The teaching of listening comprehension*. London: British Council, pp. 30-46.
- Fernández Polo, F. J. (2014) 'Native and Non-Native Speaker Interpersonal Skills at

- Conferences: Managing Self-Mentions and Humour'. In: Lyda, A. and Warchal, K. (eds) *Occupying Niches: Interculturality, Cross-culturality and Aculturality in Academic Research*. Heidelberg: Springer, pp. 163-178.
- Forey, G. and Feng, D. (2016) 'Interpersonal Meaning and Audience Engagement in Academic Presentations: A Multimodal Discourse Analysis Perspective'. In: Hyland, K. and Shaw, P. (eds) *The Routledge Handbook of English for Academic Purposes*. Oxon: Routledge, pp. 416-430.
- Fortanet-Gómez, I. (2005) 'Honoris Causa speeches: An approach to structure', *Discourse Studies* 7(1): 35-51.
- Goldman, R., Erickson, F., Lemke, J. and Derry, S. (2007) 'Selection in video'. In: Derry, S. (ed) *Guidelines For Video Research In Education: Recommendations From An Expert Panel*. Chicago: Data Research and Development Center at the University of Chicago, pp. 19-27.
- Good, J.S. (2015) 'Reported and enacted actions: Moving beyond reported speech and related concepts', *Discourse Studies* 17(6): 663-681.
- Graddol, D. (2003) 'The decline of the native speaker'. In: Anderman, G.M. and Rogers, M. (eds) *Translation today: Trends and perspectives*. London: Multilingual Matters, pp. 152-167.
- Hadar, U., Steiner, T.J., Grant, E.C. and Clifford Rose, F. (1983) 'Kinematics of Head Movements Accompanying Speech during Conversation', *Human Movement Science* 2(1-2): 35-46.
- Halliday, M. (1994) *An Introduction to Functional Grammar* (2nd ed). London: Arnold.
- Hood, S. and Forey, G. (2005) 'Introducing a conference paper: Getting interpersonal

- with your audience', *Journal of English for Academic Purposes* 4(4): 291-306.
- Hyland, K. (2009) *Academic Discourse*. London: Continuum.
- Jenkins, J. (2011) 'Accommodating (to) ELF in the international university', *Journal of pragmatics* 43: 926-936.
- Kendon, A. (2002) 'Some uses of the head shake', *Gesture* 2(2): 147-182.
- Kendon, A. (2004) *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Knapp, M. and Hall, J. (1992) *Non verbal communication in human interaction*. Texas: Holt Rinehart and Winston.
- Kress, G. (2010) *Multimodality: A Social Semiotic Approach to Contemporary Communication* (2nd ed) London: Taylor and Francis.
- Lakoff, R. (1982) 'Persuasive Discourse and Ordinary Conversation, with Examples from Advertising'. In: Tannen, D. (ed) *Analysing discourse: text and talk*. Washington: Georgetown University Press, pp. 25-42.
- McClave, E. (2000) 'Linguistic functions of head movements in the context of speech', *Journal of Pragmatics* 32: 855-878.
- McNeill, D. (1992) *Hand and mind*. Chicago: University of Chicago Press.
- Morell, T. (2015) 'International conference paper presentations: A multimodal analysis to determine effectiveness', *English for Specific Purposes* 37: 137-150.
- Morton, J. (2009) 'Genre and disciplinary competence : A case study of contextualisation in an academic speech genre', *English for Specific Purposes* 28(4): 217-229.
- Ochs, E., Gonzales, P. and Jacoby, S. (1994) 'Interpretative journeys: How physicists talk and travel through graphic space', *Configurations* (1): 151-171.

- Poyatos, F. (1983) 'Language and Nonverbal Sign Systems in the Structure of Face to Face Interaction', *Language and Communication* 3 (2):129-40.
- Querol-Julián, M. (2011) *Evaluation in discussion sessions of conference paper presentations: a multimodal approach*. Saarbrücken: LAP Lambert Academic Publishing GmbH & Co. K.G.
- Räisänen, C. (2002) 'The conference forum: A system of interrelated genres and discursive practices'. In: Ventola, E., Shalom, C. and Thompson, S. (eds) *The language of conferencing*. Frankfurt: Peter Lang, pp. 69-93.
- Rendle-Short, J. (2006) *The Academic Presentation: Situated Talk in Action*. Ashgate.
- Rowley-Jolivet, E. and Carter-Thomas, S. (2005) 'The rhetoric of conference presentation introductions: context, argument and interaction', *International Journal of Applied Linguistics* 15(1): 45-70.
- Ruiz-Garrido, M. (2015) 'Intensifying Adverbs in Academic Spoken Discourse: A Contrastive Study between English and Spanish'. In: Fortanet-Gómez, I. and Crawford Camiciottoli, B. (eds) *Multimodal Analysis in Academic Settings*. London: Routledge, pp. 61-82.
- Ruiz-Madrid, N. and Fortanet-Gómez, I. (2015) 'Contrastive Multimodal Analysis: Conference Plenary Lectures in English and in Spanish'. In Crawford-Camiciottoli, B. and Fortanet-Gómez, I. (eds.) *Multimodal Analysis in Academic Settings. From Research to Teaching*. London: Routledge, pp. 39-60.
- Ruiz-Madrid, N. and Fortanet-Gómez, I. (2015) 'A Multimodal Discourse Analysis Approach To Humour In Conference Presentations: The Case of Autobiographic References', *Procedia – Social and Behavioral Sciences* 173: 246-251

- Sidnell, J. (2006) 'Constructing gesture, talk and gaze in reenactments', *Research on Language and Social Interaction* 39(4): 377-410.
- Soulaimani, D. (2018) 'Talk, voice and gestures in reported speech: toward an integrated approach', *Discourse Studies* 20(3): 361-376.
- Swales, J. (1990) *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. and Burke, A. (2003) "'It's Really Fascinating Work': Differences in Evaluative Adjectives across Academic Registers". In: Leistyna, P. and Meyer, C.F. (eds) *Corpus Analysis: Language Structure and Language Use*. Amsterdam and Atlanta: Rodopi, pp. 1-18.
- Thompson, S. (2002) "'As the Story Unfolds': The Uses of Narrative in Research Presentations". In: Ventola, E., Shalom, C. and Thompson, S. (eds) *The language of conferencing*. New York: Peter Lang, pp. 147-183.
- Valeiras-Jurado, J. (2015) 'A Multimodal Approach to Persuasion in Conference Presentations'. In Crawford-Camicciottoli, B. and Fortanet-Gómez, I. (eds.) *Multimodal Analysis in Academic Settings. From Research to Teaching*. London: Routledge, pp. 108-130.
- Valeiras-Jurado, J. and Ruiz-Madrid, N. (2015) 'A multimodal approach to product presentations', *Procedia - Social and Behavioral Sciences* 173: 252-258.
- Valeiras-Jurado, J., Ruiz-Madrid, N. and Jacobs, G. (2018) 'Revisiting persuasion in oral academic and professional genres. Towards a methodological framework for multimodal discourse analysis of research dissemination talks', *Iberica* 35: 93-118.

Ventola, E. (1999) ‘Semiotic Spanning at Conferences; Cohesion and Coherence in and across conference papers and their discussions’. In: Bublitz, W., Lenk, U. and Ventola, E. (eds) *Coherence in Spoken and Written Discourse. How to create it and how to describe it*. Amsterdam: Benjamins, pp. 101-125.

Wulff, S., Swales, J., and Keller, K. (2009) “‘We Have about Seven Minutes for Questions’: The Discussion Sessions from a Specialized Conference”, *English for Specific Purposes* 28(2):79-92.

ⁱ <http://www.fon.hum.uva.nl/praat>

ⁱⁱ <http://tla.mpi.nl/tools/tla-tools/elan/>

ⁱⁱⁱ Tone units are sequentially listed, prominent syllables are capitalized, superscript syllables represent high key or termination, subscript syllables represent low key or termination, arrows indicate rise or fall tone when relevant for the interpretation of persuasion, pauses and their duration in seconds are indicated within brackets.